Exploring the Addition of Audio Input to Wearable Punch Recognition

Juan Quintero Ovalle
University of Bristol
Bristol, United Kingdom
jq14772@my.bristol.ac.uk

Katarzyna Stawarz
University of Bristol
Bristol, United Kingdom
k.stawarz@bristol.ac.uk

Asier Marzo
Universidad Pública de Navarra
Pamplona, Spain
asier.marzo@unavarra.es

ABSTRACT

Martial arts can promote healthy lifestyles, improve self-confidence and provide self-defence skills. Previous work has demonstrated that inertial sensors can be used to recognise movements such as punches in boxing and support self-directed training. However, many martial arts do not use gloves which means that punches can be performed with different parts of the hand, and therefore produce a different sound on impact. We investigate if it is possible to recognise different punches executed with a bare hand, and if the recognition rate improves by combining audio input with the traditional inertial sensors. We conducted a pilot study collecting a total of 600 punches, using a wearable wristband to capture inertial data and a stand-alone microphone for audio input. The results showed that it was possible to distinguish five types of punches with 94.4% accuracy when using only inertial data, and that adding audio input did not improve the accuracy. These findings can guide the design of future wearables for punch recognition.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; **Interaction devices**; • **Hardware →** *Sensors and actuators*; Sound-based input / output.

KEYWORDS

martial arts; inertial sensors; machine learning; wearables; gesture recognition; punch recognition

Interacción '19, June 25–28, 2019, Donostia-San Sebastián, Spain

ACM ISBN xxx-x-xxxx-xxxx-x/xx/xx…$15.00
https://doi.org/xx.xxxx/xxxxxxx.xxxxxxx

ACM Reference Format:
Juan Quintero Ovalle, Katarzyna Stawarz, and Asier Marzo. 2019. Exploring the Addition of Audio Input to Wearable Punch Recognition. In *The 20th International Conference on Human Computer Interaction Interacción 2019, June 25–28, 2019, Donostia-San Sebastián, Spain.* ACM, New York, NY, USA, 4 pages. https://doi.org/xx.xxxx/xxxxxxx.xxxxxxx

1 INTRODUCTION

Martial arts provide several health benefits: it is a physical activity that improves balance, the sense of psychological well-being and creates opportunities for socializing with others [15]. Additionally, it strengthens self-regulation [8] and reduces aggressive behaviours in the youth [6]. Although the chance of an injury in football or basketball is higher than in martial arts [1], martial arts training can be dangerous. The main causes of injuries are the lack of experience [1] and the type of employed technique [15]. Therefore, inexperienced martial artists who train alone are at risk. Using the adequate type of input data, a wearable device could help to support self-directed training.

Movement and punch detection

Existing devices for tracking general full-body movement are usually based on two types of technologies. On the one hand, Kinect or similar depth sensors can be used to track the body [5], but they do not perform well when detecting contact and fast movements. On the other hand, full-body tracking wearable systems exist (e.g. [9, 16]), but they are expensive and can be difficult to put on.

Martial arts involve fast movements that often finish on impact. Traditionally, these types of movements have been recognised by augmenting training equipment with small sensors, which provided more granular feedback; projects like Smart Glove [12] or an interactive scoring system for Taekwondo [3] show that this approach is feasible for both boxing and traditional martial arts. However, these systems rely on specialized sports equipment. Recent work shows that using a small wrist-based wearable with inertial sensors can recognise different types of activities [2, 11], including punches [10]. In the latter example, Minakov and Passerone showed through a small pilot study that it was possible to distinguish between two different boxing punches (cross

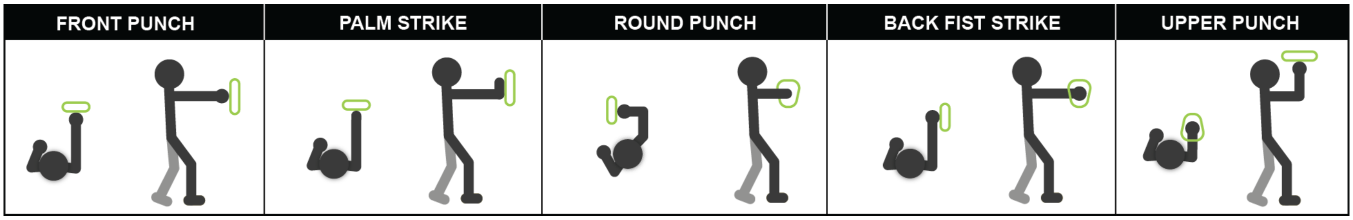

Figure 1: Types of Choi Kwang Do punches used during data collection, viewed from the top and the right.

punch and right hook), and to roughly obtain the resulting force of these punches. More recently, PIQ released a boxing hand wrap [7] that measures speed and recognises different types of punches. However, punches in boxing are performed with gloves, whereas in traditional martial arts they can be executed with bare hands. The use of bare hands also means that different techniques use different parts of the hand, e.g. the palm or the side of the fist. Distinguishing these types of punches has not been explored before.

Audio and movement recognition

While in movement recognition audio is commonly used as output, mainly to provide user feedback or instructions [13], audio input has also been used to detect the general activity of users [14]. Hitting the target with various parts of the hand can result in a different sound signature (e.g. hitting the target with a fist vs. an open hand), which suggests that audio input could potentially be used as a source of additional information to improve accuracy. Therefore, a wearable system that combines inertial sensors with audio input may provide a more accurate punch recognition and could minimise the risk of executing the punches incorrectly.

2 PILOT EVALUATION

We hypothesised that some punches would have a similar inertial signature and would be hard to differentiate using just the inertial sensors, therefore the use of sound would improve the detection rate. To test this hypothesis, we selected Choi Kwang Do (CKD) as an example martial art, because it includes a wide selection of punches [4], including several punches that do not use the front of the fist to hit the target. The training is performed without gloves which allows to practice these techniques. We identified five different types of CKD punches: front punch, palm strike, round punch, back fist strike and upper punch. Front punch and palm strike involve the same forward movement, but use a different part of the hand to hit the target. Round punch and upper punch have a similar movement, albeit rotated by 90°, so the sound signature may also be different. Back fist strike uses the top of the hand on impact, which usually produces a different type of sound compared to hitting the target with the front of the fist. All punches are illustrated in Figure 1.

Materials

Hardware. We used a simple wearable wristband to record the motion data (see Figure 2). The wristband contained an inertial sensor (i.e. three-axes accelerometers and gyroscopes; GY-521) and an Arduino Nano. The samples from the inertial sensors were captured at 200 Hz and 12 bit resolution and sent to the computer for further processing using Serial Port UART operating at 115200 bauds. To record the sound, we used a stand-alone small diaphragm, cardioid-pattern condenser microphone (SE Electronics SE1A). The microphone captured audio at 44000kHz with 16 bits resolution via an external soundcard with built-in microphone pre-amp (Behringer U-Phoria UMC204HD).

Software. The computer received the inertial data from the Arduino Nano and recorded the sound from the microphone. A program in Python applied a low-pass filter to the inertial data. The segmentation consisted of selecting an interval that represented a punch from the continuous stream of data coming from the inertial sensors and the audio from the microphone. We segmented the events by peak detection both for the inertial and the sound data. When a peak beyond a certain threshold was detected, 100ms before and 500ms after the peak were marked as an event. Figure 3 shows an

Figure 2: Top: the prototype attached to the inside of a hand. Bottom: inside of the prototype.

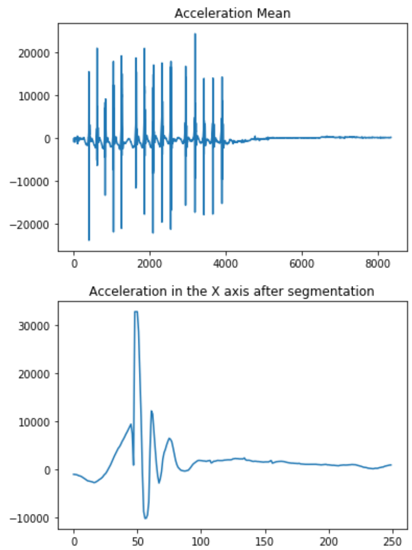

Figure 3: Top: Continuous inertial data from the x-acceleration. Bottom: A single punch event extracted from the segmentation.

example of continuous data from the sensor and a punch event extracted from the segmentation.

Various features were extracted from each event and fed into the classifier. From each axis of the accelerometer and gyroscope we obtained: mean, standard deviation, maximum, minimum and zero crossings. From the audio signal we extracted: mean, standard deviation and the three largest peaks from the Fourier Transform (frequency and phase). We used Tensor Flow's FTRL Linear Classifier with learning rate power=-0.69, learning rate=0.00001, and l1 regularization strength=0.1.

Training Data Collection and Testing

Participants. To train and test the classifier, we gathered training data from three participants: a CKD instructor (female, 35 years old) and two amateur martial artists with some kickboxing experience (both male, 23 years old).

Procedures. The amateurs with kickboxing experience were first shown how to correctly perform CKD techniques. Next, to capture the punching data, participants had to wear the wristband on their dominant hand and stand in front of the researcher who was holding two focus mitts. The microphone was placed 1.5 meters away from the punching position. While the participant was performing the punches, the researcher would adjust the position of the focus mitt to support forward, side and upward movements.

For every training round, participants were asked to perform 25 punches: each of the five punches was repeated five times, i.e. 5 front punches, 5 palm strikes, 5 round punches, 5 back fist strikes and 5 upper punches. Users performed multiple rounds of 25 punches, which resulted in 500 punches collected. The classifier was trained with this data and the accuracy was evaluated using cross-validation (20%/80% test/training set size). Additionally, 100 punches were collected to test the classifier further. Overall, we collected 600 punches: the CKD instructor performed 300 punches and amateurs performed 150 punches each. All punches were performed with the dominant hand.

3 RESULTS AND DISCUSSION

Our aim was to explore how well punches performed with the bare hand were detected using inertial sensors, and if the addition of audio input could improve the accuracy. We selected five punches from CKD as no gloves are worn during training and punches are executed using different hand positions. No previous study has investigated the use of wearables in the context of martial arts with these characteristics.

Overall, with combined inertial sensors and audio data, the classifier achieved an overall accuracy of 93%. With only inertial data, the accuracy was 94.4%. In terms of recognising individual punches, the system had no issues differentiating between round punches and back fist strikes (see Table 1). As expected, there was some confusion between front punches and palm strikes since both techniques have a similar motion. Adding the audio input from the microphone improved these specific cases but did not generally improve the results. Also, upper punches were sometimes confused with round punches and back fist strikes, the error rates were similar regardless of whether audio input was included, with an accuracy of 75%.

The ambient sound data provided marginal benefits. While it helped to recognise punches that shared the same movement, it did not increase the overall accuracy. There may be a few explanations. Firstly, the microphone placement might have played a role, as it was positioned at some distance from the target. Secondly, as some of the data was collected from amateurs, it is possible that they did not always perform the techniques correctly, which might have resulted in ambiguous data.

Future Work and Limitations

The collected data came from three martial artists: an instructor and two amateurs. While this was enough for the purpose of this pilot study, testing with more users would be necessary to generalise the results. In particular, future work should involve more professionals to ensure the correct technique is being recorded.

For study purposes, we used wired communication to transfer data, but for real-world applications wireless communication must be used. The microphone was included as a

Table 1: Confusion matrix showing the accuracy of the system for different types of punches.

	Executed punches				
Predicted punches	**Front punch**	**Palm strike**	**Round punch**	**Back fist strike**	**Upper punch**
			Inertial sensors only		
Front punch	96%	0	0	0	0
Palm strike	4%	100%	0	0	0
Round punch	0	0	100%	0	16%
Back fist strike	0	0	0	100%	8%
Upper punch	0	0	0	0	76%
			Inertial sensors and audio		
Front punch	100%	10%	0	0	0
Palm strike	0	90%	0	0	0
Round punch	0	0	100%	0	15%
Back fist strike	0	0	0	100%	10%
Upper punch	0	0	0	0	75%

separate element to ensure good quality input. If further research shows benefits of adding audio input, future versions could integrated the microphone into the prototype. As this type of hardware is already included in most smart watches, another possibility is to directly use this kind of wearables, but a study on their robustness would be required.

4 CONCLUSIONS

Our work suggests that a wearable device on the wrist can be used to distinguish five punches performed using different parts of the hand, and thus could support martial arts training without the need for specialised equipment. This is in line with previous research on punching sports using gloves. Additionally, we showed that whilst the audio input does not improve the overall accuracy in the recognition of the punches, it could be used if the focus is on distinguishing only between similar techniques. We hope that this work fosters more research into using wearables to support self-directed training of a wider range of martial arts.

REFERENCES

[1] Richard B Birrer. 1996. Trauma epidemiology in the martial arts: the results of an eighteen-year international survey. *The American Journal of Sports Medicine* 24, 6_suppl (1996), S72–S79.

[2] Saisakul Chernbumroong, Anthony S Atkins, and Hongnian Yu. 2011. Activity classification using a single wrist-worn accelerometer. In *2011 5th International Conference on Software, Knowledge Information, Industrial Management and Applications (SKIMA) Proceedings*. IEEE, 1–6.

[3] Ed H Chi, Jin Song, and Greg Corbin. 2004. Killer App of wearable computing: wireless force sensing body protectors for martial arts. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM, 277–285.

[4] KJ Choi. 2005. *Choi Kwang Do: Optimum Health, Self Defense, Personal Development.* Choi Kwang Do Martial Art International.

[5] Connsynn Chye, Mizuki Sakamoto, and Tatsuo Nakajima. 2014. An exergame for encouraging martial arts. In *International Conference on Human-Computer Interaction*. Springer, 221–232.

[6] Anna Harwood, Michal Lavidor, and Yuri Rassovsky. 2017. Reducing aggression with martial arts: A meta-analysis of child and youth studies. *Aggression and violent behavior* 34 (2017), 96–101.

[7] PIQ Sport Intelligence. [n. d.]. Advanced Sensor - System 1 PIQ ROBOT^{TM} Blue + 1 Hand wrap. https://piq.com/boxing

[8] Kimberley D Lakes and William T Hoyt. 2004. Promoting self-regulation through school-based martial arts training. *Journal of Applied Developmental Psychology* 25, 3 (2004), 283–302.

[9] Dennis Majoe, Lars Widmer, and Juerg Gutknecht. 2009. Enhanced motion interaction for multimedia applications. In *Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia*. ACM, 13–19.

[10] Ivan Minakov and Roberto Passerone. 2015. exIMUs: An Experimental Inertial Measurement Unit for Shock and Impact Detection in Sport Applications. In *International Internet of Things Summit*. Springer, 235–249.

[11] Christos Mousas. 2017. Full-body locomotion reconstruction of virtual characters using a single inertial measurement unit. *Sensors* 17, 11 (2017), 2589.

[12] Victor Xavier Navas, Jonathan Destefano, Bon Jae Koo, Eric Doty, and David Westerfeld. 2012. Smart glove. In *Systems, Applications and Technology Conference (LISAT), 2012 IEEE Long Island*. IEEE, 1–4.

[13] Nitin Sawhney and Chris Schmandt. 1998. Speaking and listening on the run: Design for wearable audio computing. In *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No. 98EX215)*. IEEE, 108–115.

[14] Johannes A Stork, Luciano Spinello, Jens Silva, and Kai O Arras. 2012. Audio-based human activity recognition using non-markovian ensemble voting. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 509–514.

[15] Thomas W Woodward. 2009. A review of the effects of martial arts practice on health. *Wisconsin Medical Journal (WMJ)* 108, 1 (2009), 40.

[16] Liang Zhang, Guido Brunnett, Katharina Petri, Marco Danneberg, Steffen Masik, Nicole Bandow, and Kerstin Witte. 2018. KaraKter: An autonomously interacting karate kumite character for VR-based training and research. *Computers & Graphics* 72 (2018), 59–69.